Robust Methods in Time Series Models with Volatility

Wendell Q. Campano
School of Statistics
University of the Philippines Diliman

Volatility in time series data is often accounted into the model by postulating a conditionally heteroskedastic variance. In-sample prediction maybe satisfactory but the out-sample prediction is usually problematic. A test for presence of volatility through a nonparametric method is proposed. An estimation procedure for the stationary part of the model by integrating block bootstrap and AR-sieve into the forward search algorithm is also provided. Simulation studies indicated high power for the nonparametric procedure in detecting local volatilities. On the other hand, the estimation method generated robust estimates of the parameters of the time series model in the presence of temporary volatility.

Keywords: block bootstrap, AR-sieve, forward search algorithm, nonparametric test, volatility

1. Introduction

Volatility is becoming a common feature of many time series data. Certain shocks not only cause fluctuations of the irregular component of the time series, but sometimes a more complicated distortion on the structure (either mean or variance-related). In economic time series for instance, financial crisis and other exogenous events have a complicated effect on both the mean and variance of the time series, for some, it is very short-term (temporary), for others, it is lingering (persistent) and even permanent for some.

A time series $\{y_t, t \in \mathrm{Z}^+\}$, is weak stationary (or simply stationary) if $\mathrm{E}(y_t)=\mu$ and $\mathrm{V}(y_t)=\sigma^2$, i.e., both mean and variance are constant over time. Enders (1995, 68) pointed out that stationarity allows the moments of the time series to be approximated by long time averages based on a single set of realizations alone. Volatility often causes fluctuations in variance, $V(y_t)=\sigma_t^2$, resulting to heteroskedasticity. A more general form of volatility is represented by conditionality upon an exogenous variable $\{x_t, t \in \mathrm{Z}^+\}$, i.e., $V(y_t/x_t)=x_t^2\sigma^2$. In the second type of volatility, also called conditional heteroskedasticity, variance of $\{y_t\}$ also depends on variance of $\{x_t\}$, so that sudden aberration in the variance of $\{x_t\}$ is magnified into the variance of $\{y_t\}$.

Aberrant observations within a time series could seriously affect identification of models that characterize the series (e.g., Atkinson and Riani, 2002). Box-Jenkins autoregressive integrated moving average (ARIMA) estimation algorithm can yield estimates with erratic behavior when there is a structural change present in the time series (Campano and Barrios, 2011). Tsay (1986) emphasized the need to detect the presence of outliers when finding possible models for the series.

Consider a time series $\{y_t\,; t=1,\ldots s\}$ that is generated by ARMA (p, q) process but for some time, this tranquil period has been disturbed and volatilities (e.g., generalized autoregressive conditional heteroskedastic or ARCH/GARCH class) occur in $\{y_t\,; t=s+1,\ldots s'\}$. Then after a relatively short period, the data generating process (DGP) returns to the original ARMA (p, q) for $\{y_t\,;\ t=s'+1,\ldots n\}$. The disturbance may recur but after every volatility period, the series will always return to the original dominating DGP, this will be referred as local volatilities.

In this paper, a residual-based nonparametric procedure for detecting local volatilities is proposed. Residual-based testing in time series data is becoming popular, one good example is given by Paparoditis and Politis (2003) in testing for unit root. The main procedure for the volatility testing is based on Jensen and Lange (2010) that consider a GARCH(1,1) model $\sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2$, in filtering the volatility processes from a wide class of DGP, see Posedel (2005) for details of GARCH(1,1) estimation.

Modeling of conditional heteroskedasticity has been thoroughly explored in the literature, models like autoregressive conditional heteroskedastic (ARCH) model, and generalized autoregressive conditional heteroskedastic (GARCH) model, and further generalizations and other extensions has been proposed. While these models aptly captured the behavior of a given realization of the time series, (Politis, 2007) noted that ARCH/GARCH models for financial time series have been criticized for their poor performance in volatility prediction. Volatility in some time series is temporary and short-lived. Estimation based of squared loss minimization is easily affected by the presence of volatility, and as a consequence, estimates are not robust. Thus, when used in prediction beyond the current realization, if that future scenario is not equivalent to the event when volatility occurred, i.e., when it reflects a more "tranquil" period, then predictions will be far from the actual values. The procedure also includes a robust estimation procedure for a time series model in the presence of volatility.

2. Volatility Modeling

The literature proposed a wide range of framework in modeling volatile behavior in time series, e.g., parametric, nonparametric, semiparametric, Bayesian, frequency domain, among others. Some of these models and general issues of estimation under volatility in time series data are presented.

Given a stationary time series, Posedel (2005) proved ergodicity and strong stationarity for the conditional variance of the process. Furthermore, Posedel (2005) concluded that high order moments of GARCH(1,1) process exist and that GARCH processes are heavy-tailed. Hall and Yao (2003) also noted that ARCH and GARCH models address the dependency of conditional second moments, an important vehicle in simplifying the modeling process. However, the difficulty in estimation for ARCH/GARCH models was linked by Hall and Yao (2003) to the heavy-tailed errors leading to the non-normal likelihood functions.

Autocorrelation leads to the non-diagonal variance-covariance matrix in time series data even with the constant mean and constant variance assumption. Relaxing the mean and variance assumptions lead at an even more complicated form of the variance-covariance matrix. Thus, in order to mitigate this complexity in a general longitudinal data, Lin and Wang (2009) factored the dependence structure in terms of the unconstrained autoregressive and scale innovation parameters in a multivariate regression to jointly model the mean and scale covariance. The simpler form of the likelihood function now facilitates estimation of parameters given the factored variance-covariance matrix.

Weide (2002) proposed a class of multivariate GARCH hat will parameterize large covariance matrices leaving a large degrees of freedom to facilitate parameter estimation. Fan et al. (2008) also proposed a multivariate volatility process based on conditionally uncorrelated components (CUC) that is a parsimonious representation for matrix-valued processes. The method generally decomposed the high dimensional problem into several lower dimensional representations. Consistency was established and a test for the presence of CUC was proposed using the bootstrap method. In a similar context, using as GARCH in mean models, Lucchetti and Rossi (2005) provided a much better testing framework for the commonly used lagrange multiplier test (actual test size), specially when the process (GARCH) exhibits persistence in volatility.

In a high-frequency data, Hafner (1998) used a class of ARCH models to observe high persistence and no significant asymmetry coefficient. With the aid of a nonparametric model, significant asymmetry of the volatility function was further validated. Still, using ARCH/GARCH models, Politis (2007) noted their poor performance in volatility prediction. Furthermore, Politis (2007) proposed a model-free (alternative to parametric model) approach based on a novel normalizing and variance-stabilizing transformation.

To mitigate the predictive ability of volatility models, Shi et al. (2000) proposed a nonlinear autoregressive model and estimated this via kernels. The predictive ability over a parametric model was established. Hurvich (2005) also considered a semiparametric estimation of the memory parameter in a model that characterizes a persistent (long memory) process.

3. Nonparametric Test for Volatility

Consider a stationary time series $\{y_t; t=1,\dots n\}$ and without loss of generality, assume that it is an autoregressive moving average or ARMA(p, q) process,

$$\phi(B)y_t = \theta(B)u_t \qquad (1)$$

where $u_t \sim IID(0,1)$

$$\phi(B)=1-\phi_1 B\dots \phi_p B^p \text{ and, } \theta(B) = 1-\theta_1 B-\dots-\theta_q B_q$$

except possibly in some time segments where local volatility may have occurred, i.e., there are localized segments with volatilities. The sequence $\{u_t\}$ at those volatility segments can be captured by the equation

$$u_t = h_t v_t \qquad (2)$$

where $v_t \sim IID(0,\sigma_v^2)$ and a sequence of volatilities $\{h_t^2\}$ which are function of the past realizations and may be stochastic.

The algorithm is divided into two stages. The first stage is a robust estimation of (1). The nonparametric estimation procedure proposed by Campano and Barrios (2011) that uses a forward search algorithm and nonparametric bootstrap methods is adopted. Simulations done using the said procedure showed that it can robustly estimate with great accuracy an ARIMA model even in the presence of structural changes provided that the time series is relatively long. The second stage involves a procedure on detecting and testing the possible existence of local volatilities within the time series.

First Stage

The time series $\{y_t; t=1,\dots,n\}$ is divided into k segments. These segments can be formed from non-overlapping or overlapping moving blocks, the choice depends on the sufficiency of the number of values (one for each segment) for the nonparametric bootstrap described in step 3.

Step 1: For each of the k segments, obtain the estimates of each parameter in Equation (1). In this step, the following matrices of estimates are obtained:

$$\hat{\varphi}_{FS} = \begin{bmatrix} \hat{\varphi}_{11} & \hat{\varphi}_{21} & \cdots & \hat{\varphi}_{p1} \\ \hat{\varphi}_{12} & \hat{\varphi}_{22} & \cdots & \hat{\varphi}_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\varphi}_{1k} & \hat{\varphi}_{2k} & \cdots & \hat{\varphi}_{pk} \end{bmatrix} \text{ and } \hat{\theta}_{FS} = \begin{bmatrix} \hat{\theta}_{11} & \hat{\theta}_{21} & \cdots & \hat{\theta}_{q1} \\ \hat{\theta}_{12} & \hat{\theta}_{22} & \cdots & \hat{\theta}_{q2} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\theta}_{1k} & \hat{\theta}_{2k} & \cdots & \hat{\theta}_{qk} \end{bmatrix}$$

Step 2: Let $\hat{\varphi}_{1FS} = \begin{bmatrix} \hat{\varphi}_{11} \\ \hat{\varphi}_{12} \\ \vdots \\ \hat{\varphi}_{1k} \end{bmatrix}, \ldots, \ \varphi_{pFS} = \begin{bmatrix} \hat{\varphi}_{p1} \\ \hat{\varphi}_{p2} \\ \vdots \\ \hat{\varphi}_{pk} \end{bmatrix}$ and $\hat{\theta}_{1FS} = \begin{bmatrix} \hat{\theta}_{11} \\ \hat{\theta}_{12} \\ \vdots \\ \hat{\theta}_{1k} \end{bmatrix}, \ldots, \ \hat{\theta}_{qFS} = \begin{bmatrix} \hat{\theta}_{q1} \\ \hat{\theta}_{q2} \\ \vdots \\ \hat{\theta}_{qk} \end{bmatrix}$ from Step 1. Note that when the segments formed are overlapping, the parameters for each segment are estimated using AR-Sieve so that each segments can be considered independent from each other.

Step 3: Obtain the estimate of each parameter in Equation (1) by performing a nonparametric bootstrap using the series of forward search estimates obtained in Step 2. That is, the bootstrap estimate of ϕ_I in (1) denoted by $\hat{\varphi}_{iBS}$ using $\hat{\varphi}_{iFS} = \begin{bmatrix} \hat{\varphi}_{i1} \\ \hat{\varphi}_{i2} \\ \vdots \\ \hat{\varphi}_{ik} \end{bmatrix}$, i=1,…, p and the bootstrap estimate of θ_j in (1) denoted by $\hat{\phi}_{jBS}$ using the series $\hat{\theta}_{jFS} = \begin{bmatrix} \hat{\theta}_{j1} \\ \hat{\theta}_{j2} \\ \vdots \\ \hat{\theta}_{jk} \end{bmatrix}$, j=1,…q.

Step 4: Substitute the parameter estimates from Step 3 to Equation (1).

Second Stage

Step 5: Using the estimated model $\hat{\phi}(B)\hat{y}_t = \hat{\theta}(B)$, compute the residuals $\hat{u}_t = y_t - \hat{y}_t, t = 1,\ldots,n$.

Step 6: Divide the residual series $\{\hat{u}_t, T = 1,\ldots,n\}$ into m segments of size l. These segments can be non-overlapping or overlapping moving blocks.

Step 7: Let $h_t^2 = w + \alpha u_{t-1}^2 + \beta h_{t-1}^2$. For each of the m segments, estimate Equation (2) by quasi-maximum likelihood estimation procedure and calculate

$$T = \hat{\alpha} + \hat{\beta} \quad \text{or} \quad \underline{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_m \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_1 + \hat{\beta}_1 \\ \hat{\alpha}_2 + \hat{\beta}_2 \\ \vdots \\ \hat{\alpha}_m + \hat{\beta}_m \end{bmatrix}$$

Step 8: Characterize the empirical distribution of T through nonparametric bootstrap. Bootstrapped values of T will be denoted by T^*. Construct a right-sided (1-γ)100% confidence interval (CI) for $\alpha+\beta$.

Step 9: Define $V_i = \begin{cases} 1 & \text{if segment } i \text{ has } \textit{volatility} \\ 0 & \text{if segment } i \text{ has } \textit{no volatility} \end{cases}$. Then, local volatilities is identified from the rule $V_i = \begin{cases} 1 & \text{if } T_i \notin CI \\ 0 & \text{otherwise} \end{cases}$, i=1,..., m.

4. Robust Estimation of Time Series

Suppose that a time series $\{y_t\}$ is represented by a general linear process

$$\Phi^p(B)(1-B^d)y_t = \Theta^q(B)u_t \tag{3}$$

Following Campano and Barrios (2011), the time series $\{y_t\}$ is cut into blocks of constant length. The forward search algorithm [see for example, Atkinson and Riani (2000), Cerioli et al., (2007), or Riani (2004), for detailed discussions of the forward search algorithm] was modified by estimating the time series in (3) for each block. When localized volatility (or conditional heteroskedasticity) is present in the time series, its effect can be contained in one or few of these blocks formed. Thus, by bootstrapping on the estimates obtained from each of the blocks, a more robust estimate of the parameters can be obtained.

The length of the block is crucial in potentially containing certain conditional heteroskedasticity into one or few blocks. Two methods of forming the blocks are proposed: mutually exclusive and overlapping blocks. Mutually exclusive block is initially desirable since each block can produce independent estimates of the parameters. However, this requires longer time series length and some persistent behavior of the time series may not necessarily be captured by a block. Block length can be adjusted but this will again require longer time series data. Overlapping blocks may also be formed. However, overlapping blocks will produce estimates that are not necessarily independent. Thus, ordinary bootstrap method will not apply. In this case, AR-sieve bootstrap can be used. Dependent estimates are replicated to characterize the empirical behavior of these estimates, see Bühlmann (1997, 2002) for more details of the AR-sieve and other bootstrap methods for time series.

5. Simulation Studies

Two simulation studies were designed to assess the performance of the nonparametric test for local volatility and the estimation of the stationary part of the time series that is affected by local volatility.

5.1 Testing for volatility

Although the simulation study presented here only covers volatilities coming from the ARCH class, the idea can be extended to other volatility processes provided that the series comes from a very high frequency data so that the convergence of $\hat{\alpha} + \hat{\beta}$ may be invoked.

To evaluate the performance of the proposed test, three autoregressive moving average DGP were considered: AR(1), MA(1), and ARMA(1,1). The local volatilities were generated from the pool of commonly used models in volatility modeling – ARCH(1) and GARCH(1,1) processes. Table 1 summarizes the details of the DGP used in the simulation.

Table 1 Specification of the DGP and Volatility Process Considered in the Simulation

Process	Model
AR(1)	$(1–0.5B)(Y_t–10)=u_t$
MA(1)	$(Y_t–10)=(1+0.65B)u_t$
ARMA(1,1)	$(1– 0.4B)(Y_t–10)=(1+0.5B)u_t$
Local Volatility	$h_t^2 = 0.1 + 0.3X_{t-1}^2 + 0.6h_{t-1}^2$
Local Volatility	$h_t^2 = 0.1 + 0.5X_{t-1}^2$

Note: u_t ~ iid (0,1)

5.2 Robust estimation of time series model with volatility

To evaluate the proposed estimation procedure, time series data under different scenarios were simulated. Parameters of the AR(1) and ARMA(1,1) models were modified to simulate a stationary and near-nonstationary process. Similarly, the parameters of MA(1) and ARMA(1,1) were modified to simulate an invertible and nearly non-invertible process. Table 2 summarized the data generating models used in the simulation.

Table 2 Models Used in the Simulation

Process	Model	Type of Process
AR(1)	$(1–0.5B)(Y_t–10) = u_t$	Stationary
	$(1–0.95B)(Y_t–10) = u_t$	near nonstationary
MA(1)	$(Y_t–10) = (1+0.5B) u_t$	Invertible
	$(Y_t–10) = (1+0.95B)u_t$	near noninvertible
ARMA(1,1)	$(1–0.4B)(Y_t–10) = (1+0.5B)u_t$	stationary and invertible
	$(1–0.95B)(Y_t–10) = (1+0.95B)u_t$	near nonstationary and near noninvertible

Note: u_t~iid (0,1) for all models

Following Lucchetti and Rossi (2005), temporary structural change was simulated from a GARCH in mean model given by:

$$Y_t = 13 + \gamma h_t + u_t \quad (4)$$

where $u_t = \sqrt{h_t} v_t$, $h_t = \kappa + \delta h_{t-1} + \alpha u_{t-1}^2$, and $v_t \sim N(0,1)$. The following values for the parameters of Model (4) were used: γ=0.5, κ=0.1, δ=0.2, , α=0.75. A total of 1,560 time point were generated for each model. The data generated from the models in Table 2 comprise the realization during a calm episode of the process. Volatility is induced by replacing a fixed number of time points at the beginning, middle and the end of the time series from the data generated from Model (4). The temporary volatility lengths were also varied.

6. Results and Discussions

The results of the simulation studies for testing volatility are presented in this section. Furthermore, an evaluation of the robustness of then estimation procedure for the DGP is also shown.

6.1 Testing the presence of volatility

Three sets of realizations from each ARMA process were generated and a local volatility model was embedded in each realization. The results for AR(1), MA(1), and ARMA(1,1) with volatilities are presented in Tables 3, 4, and 5 respectively.

It is evident from Tables 3-5 that the procedure can easily detect presence of volatility period, accuracy rate is mostly 100%. Note that the local volatilities are contained in only 20% of the segments. Increasing number of local volatility segments can decrease accuracy of detection (see Table 6).

The results came from an ideal scenario where the segment with volatility was completely isolated from those without volatility. Obviously, the exact boundaries of local volatilities are almost impossible to determine. However, by applying the *second stage* of the algorithm to overlapping moving blocks in small steps, one will have an idea on the approximate onset of the volatility period. When the overlaps of two consecutive segments are relatively large, the goal is to find consecutive segments with persistent volatility.

We consider overlapping moving blocks 10 observations at a time of the AR(1) scenario in Table 1. The procedure identifies volatility to be persistent in segments within the time points 236 to 315 when the actual volatility is introduced between time points 201 to 300. Note that the validity of using overlapping blocks in the second stage (particularly in step 8) of the procedure lies on the fact that under the null hypothesis that there are no volatilities, i.e., the error terms from an ARMA process are independent and identically distributed (IID).

Table 3 Detected Volatility Segments from AR(1) (*n*=500, *m*=5, *l*=100)

Series 1						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	5.23E-23	0	0.5141	0.9870	0
2	101-200	2.8E-05	0	0.5141	0.6850	0
3	**201-300**	**0.856832**	**0**	**0.5141**	**<.0001**	**1**
4	301-400	3.7E-05	0	0.5141	0.6790	0
5	401-500	1.13E-23	0	0.5141	1.0000	0
Series 2						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	8.00E-02	0	0.5520	0.9920	0
2	101-200	0.671444	0	0.5520	0.0010	1
3	**201-300**	**0.665539**	**0**	**0.5520**	**0.0120**	**1**
4	301-400	0.054148	0	0.5520	1.0000	0
5	401-500	1.76E-01	0	0.5520	0.9280	0
Series 3						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	8.59E-23	0	0.5278	1.0000	0
2	101-200	3.46E-06	0	0.5278	0.9800	0
3	**201-300**	**0.879664**	**0**	**0.5278**	**<.0001**	**1**
4	301-400	0.026159	0	0.5278	0.6850	0
5	401-500	1.13E-05	0	0.5278	0.9300	0

Note: Segments in bold represent true location of volatilities.

Table 4 Detected Volatility Segments from MA(1) (*n*=500, *m*=5, *l*=100)

Series 1						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	4.43E-02	0	0.4937	0.7860	0
2	101-200	5.69E-23	0	0.4937	1.0000	0
3	**201-300**	**0.808101**	**0**	**0.4937**	**<.0001**	**1**
4	301-400	6.08E-06	0	0.4937	0.9910	0
5	401-500	1.03E-01	0	0.4937	0.6700	0
Series 2						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	6.94E-02	0	0.3065	0.6830	0
2	101-200	-3.8E-23	0	0.3065	0.9990	0
3	**201-300**	**0.510838**	**0**	**0.3065**	**0.0010**	**1**
4	301-400	2.76E-06	0	0.3065	0.9920	0
5	401-500	4.08E-05	0	0.3065	0.9270	0

Series 3						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	3.65E-06	0	0.5522	1.0000	0
2	101-200	0.119011	0	0.5522	0.6630	0
3	**201-300**	**0.920395**	**0**	**0.5522**	**<.0001**	**1**
4	301-400	2.74E-05	0	0.5522	0.9180	0
5	401-500	1.88E-05	0	0.5522	0.9470	0

Note: Segments in bold represent true location of volatilities.

Table 5 Detected Volatility Segments from ARMA(1,1) (n=500, m=5, l=100)

Series 1						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	7.12E-02	0	0.3796	0.9090	0
2	101-200	0.5331	0	0.3796	<.0001	1
3	**201-300**	**0.415854**	**0**	**0.3796**	**0.0210**	**1**
4	301-400	8.07E-23	0	0.3796	1.0000	0
5	401-500	6.26E-02	0	0.3796	0.9180	0
Series 2						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	9.60E-02	0	0.4082	0.7330	0
2	101-200	6.15E-07	0	0.4082	0.9920	0
3	**201-300**	**0.680301**	**0**	**0.4082**	**0.0010**	**1**
4	301-400	4.99E-23	0	0.4082	1.0000	0
5	401-500	1.90E-01	0	0.4082	0.5120	0
Series 3						
Segment	Timepoints	T	lower CL	upper CL	P(T*>T)	Vi
1	1-100	1.68E-05	0	0.5127	0.9280	0
2	101-200	2.72E-06	0	0.5127	0.9940	0
3	**201-300**	**0.85446**	**0**	**0.5127**	**0.0010**	**1**
4	301-400	0.026603	0	0.5127	0.6870	0
5	401-500	-3.71E-23	0	0.5127	0.9990	0

Note: Segments in bold represent true location of volatilities.

Table 6 suggests that greater proportion of segments that exhibit volatility can reduce the power of the test. However, increasing the number of segments *m* can improve power but this essentially shortens the block length *l*. These results are consistent with the characteristics of procedures that use nonparametric bootstrap method.

Table 6 Power of The Test When Volatility is Present in 30% of the Segments

		Mean	Min	Max
n=1000, m=20, l=50	AR(1)	93.33	90.00	95.00
	MA(1)	93.33	85.00	100.00
	ARMA(1,1)	88.33	85.00	90.00
n=1000, m=10, l=100	AR(1)	83.33	80.00	90.00
	MA(1)	83.33	80.00	90.00
	ARMA(1,1)	83.33	80.00	90.00

6.2 Estimation under volatility

Four estimators are compared. CLS is the estimate from the conditional least squares that ignored the temporary volatility, i.e., volatility is not included into the model. The predictive ability of CLS and the proposed method are comparable. Hence, the discussion focuses on the robustness of the parameter estimates.

For the proposed method, three methods of block formation are considered. BS1: Forms overlapping segments of blocks with moving length of 12 points. BS2: Blocks are formed from overlapping segments of moving length of 60 points. BS3: Non-overlapping blocks of length 120 points each.

In a stationary AR model, robustness of the proposed estimation is clearly illustrated as the estimates produced are very close to the actual values of the parameters during the non-volatile periods. As the length of temporary volatility increases, the estimates becomes farther from the true values, but the overlapping moving blocks still provides estimates that are very near the true parameter values. For a near-nonstationary AR model, the length of volatile episodes still affect the parameter estimates, becomes farther from the true values as volatility becomes more persistent. The proposed estimation method is at its best comparable to CLS, in some cases, CLS is even better. For near-nonstationary AR model, the inferior performance of the estimator is the effect of cutting the time series into blocks. The near-nonstationarity of the series is exaggerated to appear to be non-stationary in a local setting captured by the blocks, similar to the observations (Dumanjug et al., 2010). See Tables 7 and 8 for details.

Robustness of the estimates from the proposed method in an MA model is exhibited both in Tables 9 and 10. CLS estimates becomes farther from the true values as more time points are involved in the volatile segments. Estimates from the proposed method do not deviate much from the true parameters even as more points exhibit the volatile behavior. For near non-invertible MA models, the proposed method still produced estimates that are closer to the true parameter values than what is produced by CLS.

The estimates from a stationary and invertible ARMA model are given in Table 11, while those from a near non-stationary and near non-invertible ARMA are given in Table 12. Estimates from the proposed estimation method are closer to the actual values for both the AR and MA parameters in Tables 11 and 12.

Table 7 Parameter Estimates from Data Simulated from $(1-0.5B)(Y_t-10)=a_t$.

Method	Parameters					
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\phi}$ (s.e.)	% diff	MAPE	MSE
(a) 12 points temporary volatility						
CLS	10.17	1.70	0.58	16.44	8.07	1.04
	(0.062)		(0.021)			
BS1	10.10	1.04	0.50	0.90	8.10	1.05
	(0.017)		(0.008)			
BS2	10.15	1.45	0.52	3.22	8.10	1.05
	(0.041)		(0.017)			
BS3	10.19	1.87	0.54	8.50	8.10	1.04
	(0.056)		(0.026)			
(b) 60 points temporary volatility						
CLS	10.45	4.47	0.71	42.65	8.44	1.18
	(0.096)		(0.018)			
BS1	10.26	2.60	0.51	1.58	8.59	1.29
	(0.042)		(0.013)			
BS2	10.30	2.96	0.51	2.18	8.59	1.28
	(0.107)		(0.03)			
BS3	10.38	3.84	0.54	8.79	8.57	1.25
	(0.175)		(0.043)			
(c) 120 points temporary volatility						
CLS	10.90	9.04	0.76	52.65	9.19	1.85
	(0.145)		(0.016)			
BS1	10.68	6.81	0.48	4.96	9.90	2.23
	(0.11)		(0.012)			
BS2	10.78	7.81	0.42	15.49	10.35	2.37
	(0.269)		(0.032)			
BS3	10.90	8.97	0.38	23.15	10.81	2.49
	(0.468)		(0.044)			

Note:CLS – Conditional Least Squares, BS1 – overlapping segments moving a length of 1 year, BS2 – overlapping segments moving a length of 5 years, BS3 – non-overlapping 10-year block

Table 8 Parameter Estimates from Data Simulated from $(1-0.95B)(Y_t-10)=a_t$.

Method	Parameters					
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\phi}$ (s.e.)	% diff	MAPE	MSE
(a) 12 points temporary volatility						
CLS	10.69	6.90	0.94	1.41	10.25	1.61
	(0.473)		(0.009)			
BS1	4.17	58.32	0.92	3.00	10.45	1.89
	(3.014)		(0.006)			
BS2	6.51	34.92	0.92	3.13	10.18	1.72
	(3.371)		(0.013)			
BS3	10.61	6.13	0.94	0.82	10.19	1.61
	(0.857)		(0.014)			
(b) 60 points temporary volatility						
CLS	10.52	5.25	0.91	4.48	9.63	1.53
	(0.328)		(0.011)			
BS1	10.07	0.66	0.89	5.94	9.63	1.54
	(0.209)		(0.006)			
BS2	10.32	3.25	0.89	6.47	9.70	1.54
	(0.397)		(0.015)			
BS3	9.97	0.29	0.89	6.73	9.66	1.54
	(0.608)		(0.024)			
(c) 120 points temporary volatility						
CLS	11.92	19.19	0.92	2.78	10.80	1.92
	(0.435)		(0.01)			
BS1	12.12	21.16	0.87	8.80	11.83	1.96
	(0.337)		(0.016)			
BS2	12.20	22.01	0.82	13.29	13.02	2.05
	(0.88)		(0.051)			
BS3	11.01	10.08	0.80	16.29	13.10	2.15
	(1.266)		(0.101)			

Note: CLS – Conditional Least Squares, BS1 – overlapping segments moving a length of 1 year, BS2 – overlapping segments moving a length of 5 years, BS3 – non-overlapping 10-year block

Table 9 Parameter Estimates from Data Simulated from $(Y_t - 10) = (1+0.5B)\, a_t$

Method	Series 3					
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\theta}$ (s.e.)	% diff	MAPE	MSE
(a) 12 points temporary volatility						
CLS	10.09	0.88	-0.47	6.96	8.53	1.22
	(0.041)		(0.022)			
BS1	10.04	0.41	-0.48	3.42	8.50	1.22
	(0.012)		(0.006)			
BS2	10.06	0.63	-0.48	3.96	8.51	1.22
	(0.029)		(0.014)			
BS3	10.09	0.88	-0.48	4.04	8.53	1.22
	(0.039)		(0.022)			
(b) 60 points temporary volatility						
CLS	10.46	4.59	-0.59	18.66	9.57	1.74
	(0.053)		(0.02)			
BS1	10.27	2.70	-0.51	1.35	9.45	1.78
	(0.058)		(0.006)			
BS2	10.33	3.30	-0.51	1.23	9.50	1.77
	(0.138)		(0.013)			
BS3	10.45	4.54	-0.51	2.24	9.62	1.76
	(0.243)		(0.022)			
(c) 120 points temporary volatility						
CLS	10.86	8.59	-0.58	15.62	10.49	2.33
	(0.061)		(0.021)			
BS1	10.65	6.53	-0.47	5.83	10.42	2.41
	(0.097)		(0.01)			
BS2	10.74	7.44	-0.44	12.47	10.66	2.43
	(0.259)		(0.035)			
BS3	10.85	8.53	-0.37	25.25	11.13	2.52
	(0.443)		(0.055)			

Note: CLS – Conditional Least Squares, BS1 – overlapping segments moving a length of 1 year, BS2 – overlapping segments moving a length of 5 years, BS3 – non-overlapping 10-year block

Table10 Parameter Estimates from Data Simulated from $(Y_t - 10) = (1+0.95B)a_t$

Method	Series 1					
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\theta}$ (s.e.)	% diff	MAPE	MSE
(a) 12 points temporary volatility						
CLS	10.07	0.73	-0.83	12.26	8.89	1.29
	(0.053)		(0.014)			
BS1	10.03	0.29	-0.88	7.78	8.80	1.30
	(0.022)		(0.006)			
BS2	10.06	0.57	-0.87	8.80	8.83	1.29
	(0.041)		(0.012)			
BS3	10.09	0.91	-0.87	8.39	8.85	1.30
	(0.054)		(0.021)			
(b) 60 points temporary volatility						
CLS	10.40	3.99	-0.72	24.69	9.73	1.74
	(0.057)		(0.018)			
BS1	10.21	2.12	-0.84	12.10	9.38	1.84
	(0.055)		(0.011)			
BS2	10.27	2.74	-0.83	13.07	9.42	1.82
	(0.132)		(0.026)			
BS3	10.41	4.10	-0.84	11.41	9.54	1.85
	(0.212)		(0.039)			
(c) 120 points temporary volatility						
CLS	10.93	9.35	-0.61	35.98	11.26	2.71
	(0.067)		(0.02)			
BS1	10.71	7.07	-0.78	18.22	10.71	3.07
	(0.105)		(0.021)			
BS2	10.81	8.13	-0.74	21.87	10.81	2.90
	(0.27)		(0.062)			
BS3	10.92	9.20	-0.72	24.60	10.96	2.82
	(0.458)		(0.109)			

Note: CLS – Conditional Least Squares, BS1 – overlapping segments moving a length of 1 year, BS2 – overlapping segments moving a length of 5 years, BS3 – non-overlapping 10-year block

Table 11 Parameter Estimates from Data Simulated from (1–0.4*B*)(Y_t–10)=(1+0.5)*a*

Method	Series 2						MAPE	MSE
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\phi}$ (s.e.)	% diff	$\hat{\theta}$ (s.e.)	% diff		
(a) 12 points temporary volatility								
CLS	10.09	0.86	0.51	27.90	-0.38	23.52	8.56	1.16
	(0.077)		(0.029)		(0.031)			
BS1	10.04	0.40	0.39	2.23	-0.48	4.40	8.57	1.17
	(0.021)		(0.011)		(0.017)			
BS2	10.07	0.66	0.40	0.87	-0.47	6.13	8.58	1.17
	(0.047)		(0.029)		(0.036)			
BS3	10.09	0.92	0.46	15.44	-0.44	11.92	8.55	1.17
	(0.059)		(0.041)		(0.054)			
(b) 60 points temporary volatility								
CLS	10.51	5.11	0.80	101.19	0.09	117.56	9.52	1.72
	(0.154)		(0.02)		(0.033)			
BS1	10.37	3.70	0.45	11.77	-0.32	35.16	9.42	1.89
	(0.169)		(0.021)		(0.034)			
BS2	10.20	2.02	0.45	13.48	-0.32	36.71	9.34	1.89
	(0.217)		(0.052)		(0.075)			
BS3	10.03	0.26	0.52	29.08	-0.25	50.79	9.26	1.86
	(0.253)		(0.079)		(0.119)			
(c) 120 points temporary volatility								
CLS	11.29	12.90	0.93	133.46	0.47	194.53	10.29	2.23
	(0.296)		(0.011)		(0.027)			
BS1	11.67	16.70	0.51	26.83	-0.16	67.58	11.38	2.75
	(0.376)		(0.024)		(0.041)			
BS2	11.21	12.14	0.46	14.43	-0.19	61.47	10.93	2.77
	(0.518)		(0.051)		(0.094)			
BS3	11.14	11.41	0.54	34.66	-0.15	70.31	10.56	2.64
	(0.576)		(0.071)		(0.151)			

Note: CLS – Conditional Least Squares, BS1 – overlapping segments moving a length of 1 year, BS2 – overlapping segments moving a length of 5 years, BS3 – non-overlapping 10-year block

Table 12 Parameter Estimates from Data Simulated from $(1–0.95B)(Y_t–10)=(1+0.95)a$

Method	Series 1						MAPE	MSE
	$\hat{\mu}$ (s.e.)	% diff	$\hat{\phi}$ (s.e.)	% diff	$\hat{\theta}$ (s.e.)	% diff		
(a) 12 points temporary volatility								
CLS	13.05	30.51	0.94	0.84	-0.70	26.74	94.26	1.43
	(0.64)		(0.009)		(0.018)			
BS1	13.44	34.37	0.95	0.28	-0.94	0.76	98.61	1.81
	(1.568)		(0.002)		(0.011)			
BS2	10.53	5.31	0.95	0.06	-0.90	5.54	94.39	1.60
	(1.017)		(0.004)		(0.04)			
BS3	10.83	8.34	0.95	0.17	-0.84	11.53	94.76	1.50
	(1.398)		(0.01)		(0.074)			
(b) 60 points temporary volatility								
CLS	12.33	23.32	0.95	0.41	-0.12	87.60	22.47	2.62
	(0.854)		(0.008)		(0.026)			
BS1	13.71	37.08	0.95	0.00	-0.76	19.83	35.53	5.45
	(1.841)		(0.003)		(0.041)			
BS2	11.06	10.58	0.95	0.30	-0.74	22.63	33.76	5.05
	(0.881)		(0.007)		(0.096)			
BS3	11.10	11.00	0.94	1.05	-0.70	25.92	31.74	4.60
	(1.127)		(0.014)		(0.143)			
(c) 120 points temporary volatility								
CLS	10.45	4.48	0.95	0.29	-0.09	90.23	54.42	2.37
	(0.782)		(0.008)		(0.026)			
BS1	8.48	15.24	0.93	2.25	-0.64	32.38	46.09	3.71
	(0.694)		(0.012)		(0.046)			
BS2	9.60	4.05	0.87	8.00	-0.65	31.06	41.35	3.70
	(1.111)		(0.047)		(0.105)			
BS3	10.12	1.20	0.82	13.54	-0.67	29.74	52.97	3.77
	(1.597)		(0.115)		(0.175)			

6. Conclusions

Simulation studies exhibited high power of the test procedure when proportion of segments with local volatilities is relatively small. When there is a high proportion of local volatilities in the data, increasing the number of test segments can improve performance of the algorithm as long as there are enough time points within each segment for volatility model estimation.

In the presence of temporary volatility, a hybrid of the modified forward search algorithm and the bootstrap method can produce robust estimates of the parameters of the model during non-volatile episodes. Overlapping blocks are better than the non-overlapping blocks of time series in replicating the time series for re-sampling. Overlapping blocks can also be advantageous in short time series. Estimates from this method can be used in predicting out-of-sample points in the time series, conditional on the non-occurrence of volatility at those points. This will address the issue of poor prediction among volatility models noted in the literature.

References

ATKINSON, A., and M. RIANI, 2000, *Robust Diagnostic Regression Analysis*, New York: Springer-Verlag.

ATKINSON, A.C., M. RIANI, 2002, Forward search added-variable t-tests and the effect of masked outliers on model selection, *Biometrika*, 89, 939-946.

CAMPANO, W.Q. and E.B. BARRIOS, 2011, Robust estimation of a time series model with structural change, *Journal of Statistical Computation and Simulation*, 81: 7, 909 – 927.

CERIOLI, A., M. RIANI, and A. ATKINSON, 2007, Clustering contiguous or categorical data with the forward search, Bulletin of the 56th Meeting of the International Statistical Institute.

BÜHLMANN, P., 1997, Sieve bootstrap for time series, *Bernoulli*, 3:123-148.

BÜHLMANN, P., 2002, Boostraps for time series, *Statistical Science*, 17: 52-72.

DUMANJUG, C., E. BARRIOS, and J. LANSANGAN, 2010, Bootstrap procedures in a spatial-temporal model, *Journal of Statistical Computing and Simulation*, 80(7): 809-822.

ENDERS, W., 1995, *Applied Econometric Time Series*, New York: John Wiley and Sons.

FAN, J., M. WANG, Q. YAO, 2008, Modelling multivariate volatilities via conditionally uncorrelated components, *Journal of the Royal Statistical Society*, Ser. B. 70(4):679-702.

HAFNER, C., 1998, Estimating high-frequency foreign exchange rate volatility with nonparametric ARCH models, *Journal of Statistical Planning and Inference*, 68:247-269.

HALL, P., and Q. YAO, 2003, Inference in ARCh and GARCH models with heavy-tailed errors, *Econometrica*, 71(1):285-317.

HURVICH, C., E. MOULINES, and P. SOULIER, 2005, Estimating long memory in volatility, *Econometrica*, 73(4):1283-1328.

JENSEN, A.T. and T. LANGE, 2010, On convergence of the QMLE for misspecified GARCH Models, *Journal of Time Series Econometrics*, Vol. 2: Iss. 1, Article 3.

LIN, T., and Y. WANG, 2009, A robust approach to joint modeling of mean and scale covariance for longitudinal data, *Journal of Statistical Planning and Inference*, 139:3013-3026.

LUCCHETTI, R., E. ROSSI, 2005, Artificial regression testing in the GARCH-in-mean model, *Econometrics Journal*, 8:306-322.

PAPARODITIS, E., D.N. POLITIS, 2003, Residual-based block bootstrap for unit root testing, *Econometrica,* Vol. 71, No. 3, 813-855.

POLITIS, D., 2007, Model-Free Versus Model-Based Volatility Prediction, *Journal of Financial Econometrics*, 5(3):358-389.

POSEDEL, P., 2005, Properties and estimation of GARCH(1,1) model, *Metodološki zvezki*, 2(2):243-257.

RIANI, M., 2004, Estensions of the forward search to time series, *Linear and Nonlinear Dynamics in Time Series*, 8 (Article 2).

SHI, Z., Y. TAMURA, and T. OZAKI, 2000, Empirical evaluation of non-parametric autoregressive model for dynamics reconstruction and prediction, *Journal of Signal Processing*, 4(2):185-193.

TSAY, R., 1986, Time series model specification in the presence of outliers, *Journal of the American Statistical Association*, 81, 132-141.

WEIDE, R., 2002, Go-GARCH: A Multivariate generalized orthogonal GARCH Model, *Journal of Applied Econometrics*, 17: 549-564.